Detailed assessment of low-back loads may not be worth the effort: A comparison of two methods for exposure-outcome assessment of low-back pain

Coenen, P.; Kingma, I.; Boot, C.R.L.; Bongers, P.M.; van Dieen, J.H.

published in
Applied Ergonomics
2015

DOI (link to publisher)
10.1016/j.apergo.2015.06.005

Applied Ergonomics 51 (2015) 322–330

Contents lists available at ScienceDirect

Applied Ergonomics

journal homepage: www.elsevier.com/locate/apergo

Detailed assessment of low-back loads may not be worth the effort: A comparison of two methods for exposure-outcome assessment of low-back pain

Pieter Coenen [a, b, c], Idsart Kingma [b, c], Cécile R.L. Boot [c, d], Paulien M. Bongers [c, e], Jaap H. van Dieën [b, c, f, *]

[a] School of Physiotherapy and Exercise Science, Curtin University, Perth, WA, Australia
[b] MOVE Research Institute Amsterdam, Faculty of Human Movement Sciences, VU University Amsterdam, Amsterdam, The Netherlands
[c] Body@Work, Research Center on Physical Activity, Work and Health, The Netherlands
[d] Department of Public and Occupational Health, EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, The Netherlands
[e] TNO Healthy Living, Hoofddorp, The Netherlands
[f] King Abdulaziz University, Jeddah, Saudi Arabia

ARTICLE INFO

Article history:
Received 16 April 2014
Received in revised form
12 April 2015
Accepted 4 June 2015
Available online xxx

Keywords:
Low-back load
Exposure assessment
Low-back pain
Accuracy

ABSTRACT

The trade-off between feasibility and accuracy of measurements of physical exposure at the workplace has often been discussed, but is unsufficiently understood. We therefore explored the effect of two low-back loading measurement tools with different accuracies on exposure estimates and their associations with low-back pain (LBP).

Low-back moments of 93 workers were obtained using two methods: a moderately accurate observation-based method and a relatively more accurate video-analysis method. Group-based exposure metrics were assigned to a total of 1131 workers who reported on their LBP status during three follow-up years. The two methods were compared regarding individual and group-based moments and their predictive value for LBP.

Differences between the two methods for peak moments were high at the individual level and remained substantial at group level. For cumulative moments, differences between the two methods were attenuated as random inaccuracies cancelled out. Peak moments were not predictive for LBP in any method while cumulative moments were, suggesting comparable predictive values of the two methods. While assessment of low-back load improves from investing in collecting relatively more accurate individual-based data, this does not necessarily lead to better predictive values on a group level, especially not for cumulative loads.

1. Introduction

Exposure to physical risk factors at the workplace such as lifting, pushing, pulling, and awkward trunk postures (e.g., flexion and rotation) has been associated with low-back pain (LBP; da Costa et al., 2010; Griffith et al., 2012; Lötters et al., 2003). However, it has also been argued that evidence concerning such work related risk factors for LBP is weak and inconsistent (Bakker et al., 2009; Kwon et al., 2011), potentially due to insufficient high quality studies using accurate objective measurement methods (Burdorf, 2010; David, 2005). An important potential reason for this is that the choice for a measurement method for occupational physical exposure involves a trade-off between accuracy and feasibility (i.e., in time and costs). As an example, although self-reports of physical exposure are frequently used as they can be obtained with relative ease and few expenses, outcomes are highly subjective and often based on rough categorization, thereby limiting accuracy (Balogh et al., 2004; Punnett, 2004). As a result, in theory, the choice of such methods in view of available resources, is expected to affect accuracy of exposure estimates which may bias risk associations (Tielemans et al., 1998) and reduce statistical power (Mathiassen et al., 2002, 2010). However, in practice, this is not always the case in epidemiological literature, since studies that measure more

* Corresponding author. Faculty of Human Movement Sciences, van der Boechorststraat 9, 1081 BT Amsterdam, The Netherlands.
E-mail address: j.vandieen@fbw.vu.nl (J.H. van Dieën).

http://dx.doi.org/10.1016/j.apergo.2015.06.005
0003-6870/

accurately often measure limited amounts of subjects which reduces the power of the given study (Griffith et al., 2012). Therefore, the effect of the accuracy of a chosen measurement method on exposure-outcome associations for occupational physical exposure risk factors of LBP is not well understood.

Mechanical low-back load as a result of exposure to physical load at the workplace (e.g., lifting and trunk flexion) is an appropriate load measure and is expected to be an important determinant of LBP (Chaffin, 2009; Wells et al., 2004). Such loads (i.e., low-back moments or forces on the lumbar spine) are suspected to provide a direct relationship with spinal failure and consequently with LBP. Mechanical low-back loads can be obtained from measured hand forces and structured posture observations as inputs in a biomechanical model in epidemiological studies (e.g.; Neumann et al., 2001). It has however been shown that these methods can lead to large inaccuracies (de Looze et al., 1994). Nevertheless, such estimates are predictive for LBP (Coenen et al., 2013b; Norman et al., 1998) and more valid and reliable methods for mechanical low-back loads, such as direct measurement techniques (i.e., combining information from motion tracking systems and external force measurements; Kingma et al., 2010; Marras et al., 2010a; Plamondon et al., 1996) can potentially lead to more accurate estimates and to less biased associations with LBP. However, in accordance with abovementioned trade-off, such methods are often costly and difficult to apply to a field setting, as they may interfere with the work performed (Trask et al., 2007). Posture fitting on planar video recordings during manual materials handling tasks has been shown to be a feasible and accurate method for application in field settings (Chang et al., 2003; Coenen et al., 2011; Xu et al., 2011). Yet, such methods are time-consuming and only allow analysis of selected tasks rather than continuous monitoring.

To date, the optimum of the above-mentioned trade-off, indicating which measurement tool for occupational low-back load assessment should be chosen in order to have the best combination of measurement accuracy and feasibility, is unknown. We therefore explored this trade-off by comparing the assessment of low-back loads based on observations (Coenen et al., 2013b) and low-back loads assessed more accurately using detailed video-analysis (Coenen et al., 2011). The two methods were compared, at both the individual and group levels, in terms of the accuracy of load estimates and predictive values regarding LBP prevalence.

2. Methods

2.1. Population and data collection

Data were collected as part of the SMASH study (Ariëns et al., 2001; Hoogendoorn et al., 2000) involving workers in a baseline measurement protocol, in which occupational low-back load was assessed at the workplace. Workers were recruited from 34 companies in the Netherlands representing several industrial and service branches, including metal, computer software, chemical, pharmaceutical, food and wood construction industries, as well as insurance companies, childcare centers, hospitals, distribution companies and road worker organizations. The study population thus included workers performing diverse tasks with a wide range of physical and mental workloads.

During the SMASH study, videos were collected at four randomly chosen instants in one day. Videos were collected for 5–15 min during each of the four occasions, depending on the variability of the worker's task, to obtain a representative sample of the worker's jobs. During these periods, external forces at the hands were measured when present, using force transducers (for pushing and pulling tasks) or weighting scales (for lifting tasks). For pushing and pulling tasks, a measured horizontal direction of the force was assumed and a single measured value of the transducer was used. For lifting, measured weights were used as input in the two methods, as will be outlined in detail later.

A three year annual follow-up assessment of LBP was performed using a self-administered Dutch version of the Nordic Questionnaire (Kuorinka et al., 1987). LBP was defined when a worker reported regular or prolonged LBP during at least one of the three years of follow-up. This definition of LBP prevalence was independent from LBP status at baseline. Regular or prolonged LBP was assessed based on self-reports and was thus not based on medical diagnosis, nor was it related to a specific incident or cause.

For the current study, of the 1802 workers who completed the baseline questionnaires (regarding personal information such as age, gender and LBP prevalence), LBP data in at least one of the years of follow-up were available for 1131 of them. These workers were a-priori allocated to occupational groups with similar tasks and physical loads based on the International Standard Classification of Occupations. These occupations were then again, based on expert judgments composed into 23 groups. These groups, such as a group of workers performing mainly *sitting tasks with varying postures* or *alternating standing, walking and/or sitting without external forces* were solely based on the expected physical work load without any prior knowledge on the actual quantified physical work load, baseline LBP status and/or psychosocial or workplace factors. This expected physical work load was subjectively assessed after watching the video by observers that were recruited among students of the Faculty of Human Movement Sciences of the VU University, Amsterdam and were extensively trained on the task. This classification scheme has been shown to be effective, leading to substantial between-group variation in low-back load variables in earlier work (Coenen et al., 2014b). Moreover, applying a group-based measurement approach has been shown to be an efficient strategy leading to more reliable estimates of exposure, since random measurement errors in individual estimates of exposure may decrease (Hoozemans et al., 2001; Jansen et al., 2003).

For the current study, data of those 19 groups of which video material was available for at least 4 workers were used (Table 1). Videos of these workers were observed during which manual material handling (MMH) tasks, i.e., lifting, pushing and pulling, were identified. From each group, four or if available five workers were randomly selected from whom all MMH tasks that occurred during the video recording were identified. As a result, 4872 MMH tasks of a total of 93 workers were analysed in the current study (Table 1). The use of this selection has been shown to be effective in assessing exposure-outcome associations (Coenen et al., 2014a) while it has also been shown that adding more workers per group does not lead to a considerably higher precision and power of the study outcomes (Coenen et al., 2014b). Low-back moments of all identified MMH tasks were subsequently assessed using two methods that will be described in the following paragraph.

2.2. Assessment of low-back moment

All selected videos of MMH tasks were used for low-back moment assessment with two different methods that have been described in more detail previously (Coenen et al., 2014a, 2013b). In the first assessment method, a procedure was performed in which postural observation data were used as inputs to a biomechanical model (Coenen et al., 2013b). Structured continuous observations of body segment positions (i.e., trunk flexion, trunk rotation and arm elevation in the dominant arm) were applied to the complete video material. Subsequently, to get a fair comparison with the second method that will be described below, only observations of the 4872 MMH tasks of the 93 selected workers were selected for

Table 1
Data set characteristics. In the upper part of the table, the total number of workers and the number of workers observed are shown for each group. In the lower part, descriptive statistics of the workers who were included in the cohort (left column) and the workers from whom low-back moment data are available (right columns) are presented. Age, gender and LBP prevalence at baseline are shown.

Group description		Total	Analysed
Mainly sitting			
1.	Sitting with varying postures	133	5
2.	Sitting with little varying postures (computer work)	57	5
3.	Sitting with little varying postures, in awkward postures (no computer work)	31	5
4.	Sitting with little varying postures, with repetitive movements	95	5
Mainly standing work			
5.	Standing with varying postures (including walking) without external forces	26	5
6.	Standing with varying postures and small external forces	69	5
7.	Standing with varying postures and moderate external forces	87	5
8.	Standing with varying postures and large external forces	65	5
9.	Standing with varying, awkward postures and moderate external forces	66	5
Awkward postures (mainly static exposure)			
10.	Standing in static awkward posture without external forces	42	5
11.	Standing in static awkward posture with small external forces	70	5
12.	Mainly static back exposures by alternating awkward postures	28	4
Alternating exposures (standing, walking and/or sitting)			
13.	Alternating standing, walking and/or sitting without external forces	167	5
14.	Alternating standing, walking and/or sitting with small external forces	36	5
15.	Alternating standing, walking and/or sitting with moderate external forces	52	5
16.	Alternating standing, walking and/or sitting with large external forces	21	4
17.	Alternating standing and walking in static awkward postures, external forces	27	5
18	Alternating standing and walking in postures, moderate external forces	36	5
Combined functions (as a result of changes in tasks)			
19.	Combined exposures	23	5
Population descriptive			
Number of workers (n)		1131	93
Age (years)		36(9)	36(9)
Males (n (%))		800 (71%)	61 (66%)
LBP at baseline (n (%))		399 (35%)	28 (30%)

further analysis. Consequently, non MMH tasks were discarded. These data were used as kinematics input together with segment anthropometrics to construct a manikin consisting of a trunk/head, upper arms and a lower arms/hands segment (i.e., assuming symmetry in the two arms). Segment mass, length, position of the center of mass and inertia tensor were estimated based on regression equations using total body mass and stature of each worker (Zatsiorsky, 2002). Observational categories are shown in Table 2 (e.g., ‘mild trunk flexion’ depicting 30–60° trunk flexion that was assigned a value 45). Manikin kinematics was used together with a single unilateral hand force measure as input to a three-dimensional linked segment model to calculate time series of static moments at the level of L5S1. This hand force was implemented as a horizontal force for pushing/pulling tasks and as a mass with its centre of gravity between the two hands of the manikin for lifting tasks.

The second assessment method is a detailed video-analysis of the same footage of the MMH tasks (i.e., video material of the 4872 MMH tasks of the selection of workers, n = 93). In short, in a graphical user interface, a semi three-dimensional manikin, consisting of nine body segments, was constructed over four key video frames (i.e., the begin and the end frame of the tasks as well as two equally spaced frames between begin and end frames; Fig. 1). A cubic spline interpolation of the segment angles over the four key frames was applied to estimate segment angles time series over the entire MMH trajectory according to earlier work (Xu et al., 2010). This kinematics was used together with segment anthropometrics (in accordance with the first method as described above) to construct a manikin. Kinematics of this manikin was used together with segment anthropometrics and a single unilateral hand force measure, as inputs to a three-dimensional linked-segment model to calculate time series of dynamic L5S1 moments. This hand force was implemented as a horizontal force for pushing/pulling tasks and as a mass with its centre of gravity between the two hands of the manikin for a lifting task. In contrast to the first method, which uses a static approach, a dynamic model was used and input

Table 2
Observational categories. The table shows a description and corresponding values for the observed variables. The last column shows body orientation values that were used for the calculation of low-back moments.

Variable	Observation		Moment calculation
	Description	Category	Values
Trunk flexion (sagittal plane)	Neutral	<30°	0°
	Mild flexion	30–60°	45°
	Extreme flexion	60–90°	75°
	Very extreme flexion	>90°	90°
Trunk rotation (transverse plane)	Neutral	<30°	0°
	Twisting	>30°	30°
Arm elevation (sagittal plane)	Neutral	<30°	15°
	Mild elevation	30–60°	45°
	Extreme elevation	60–90°	75°
	Very extreme elevation	>90°	90°

Fig. 1. Video-analysis method. The graphical user interface depicting a three-dimensional manikin plotted onto a video frame is shown (upper part of the figure). In the lower part of the figure, a typical example of four key video frames of a field-based lifting task are shown. Beginning and ending frames of the task were selected from the video fragments by the rater. Two intermediate frames, equally spaced in time between the beginning and end frame, were automatically selected to obtain four video frames. In these four video frames, a semi three-dimensional manikin was constructed consisting of nine segments (right foot, lower leg and upper leg; pelvis, trunk/head, two upper arms, two forearms/hands). This manikin allows for semi three-dimensional analysis of movements (ankle flexion/extension, knee flexion/extension, hip flexion/extension, trunk flexion/extension, trunk rotation, trunk lateral flexion, shoulder flexion/extension, shoulder abduction and elbow flexion/extension). Furthermore, the manikin can be scaled, rotated around its longitudinal axis (axial rotation) and translated horizontally and vertically along the video frame.

segment angles were obtained from a fitted and interpolated manikin (providing more detailed information about postures rather than from categorical observation-based postures). The video-analysis method used has been shown to have good validity (Coenen et al., 2011) and inter-rater reliability (Coenen et al., 2013a). No systematic differences and a strong correspondence (correlation >0.85) of the video-analysis method compared to a laboratory based gold standard reference method (i.e., motion analysis and force plate measurements) was reported (Coenen et al., 2011). Also, there was excellent agreement among raters (i.e., intra-class correlation coefficient >0.9), while inter-rater variation was relatively low (<10 Nm; Coenen et al., 2013a).

2.3. Data analysis

Time series of the resultant of the three-dimensional low-back moment of all subjects during all MMH tasks were obtained for the two assessment methods. These time series were used to calculate peak and cumulative moments of both assessment methods. Individual peak moments were defined as the maximum moment obtained for an individual over all MMH tasks. Individual cumulative moments were assessed by calculating the area under the moment curves of all MMH tasks after which the moments were extrapolated to a work week (using the duration of the observation and the number of working hours per week). Besides these individual moments, peak and cumulative loads were also calculated at a group level in each of the exposure groups, by calculating the group mean over all individuals from whom video analyses were performed. For both assessment methods, these group values were assigned to all group members (i.e., all 1131 workers who reported on their LBP status during the three years of follow-up).

Agreement between the two ways of calculating low-back moments was evaluated by calculating intra-class correlation coefficients (ICCs), for the individual and the group-based low-back moments. ICCs <0.40 were assumed poor, ICCs 0.40–0.75 were assumed good and ICCs >0.75 were assumed excellent (Fleiss, 1986). Furthermore, data were visualized by plotting the measurement differences (errors) of the two assessment methods against their respective means, i.e., a Bland–Altman plot (Bland and Altman, 1986). In these plots, systematic and random differences between the two methods of assessing low-back moments were shown.

To assess predictive values of all metrics, univariate associations with LBP were estimated by calculating odds ratios (ORs), 95% confidence intervals and p-values using logistic regression. In these analyses, low-back moments were the (continuous, expressed per unit of low-back moments, Nm for peak moments and Nm/week for cumulative moments) independent variable and LBP the (dichotomous, either case or control) dependent variable. To facilitate the interpretation of the ORs, cumulative low-back moments were divided by 10^5. All data processing was applied using custom Matlab software (version R2011a) while statistical analyses were performed using the statistics toolbox.

3. Results

A poor correlation between estimates obtained with the two assessment methods was found (ICC = 0.28) for peak moments estimated at the individual level (Fig. 2). However, when peak moments were estimated at a group level, excellent agreement (ICC = 0.82) between the two methods was found. As can be seen in the Bland–Altman plot, for peak moments, differences between the two methods could be quite substantial at an individual level (i.e., up to 300 Nm, dashed lines in Fig. 3) with a mean systematic difference of about 40 Nm (dotted line in Fig. 3). At a group level these inaccuracies decreased substantially but were still present (i.e., up to about 50 Nm with a mean systematic difference of about 30 Nm). Especially for the moments at an individual level, differences might be proportional (i.e., differences between methods were proportionally larger at higher absolute moments).

For cumulative moments, an excellent agreement of the two methods was shown for moments at an individual level and at a group level (ICC = 0.95 and ICC = 0.94, respectively; Fig. 2). It should however be noted that these ICCs are slightly lower when the evidently outlying group was removed (ICC = 0.89 and ICC = 0.84, respectively). For cumulative moments, differences between the two methods were comparable at the individual level compared to at group level (with mean systematic difference around $2 \cdot 10^5$ Nm/week; Fig. 3).

Peak moments were not significantly associated with LBP (with ORs of 1.001(1.000–1.002) and 1.001(0.999–1.003), for the observation-based method and the video-analysis method, respectively, Table 3). Cumulative moments were significantly associated with LBP (with ORs of 1.086 (1.036–1.138) and 1.056 (1.024–1.090), for the observation-based method and the video-analysis method, respectively). Note however, that these ORs are expressed per unit of the low-back moment (i.e., Nm and 10^5 Nm/week for peak and cumulative moments, respectively). A better interpretation of these OR can be made when expressing these ORs in the difference in low-back moments corresponding with a difference of the groups with the highest mechanical load compared with the group with the lowest mechanical load. These ORs are 1.26 and 1.31 for peak low-back moments (for the observation-based method and the video-analysis method, respectively) and 6.41 and 7.25 for cumulative low-back moments (for the observation-based method and the video-analysis method, respectively). Both estimates showed comparable associations with LBP, though the more accurate moments obtained by video-analysis yielded a slightly higher OR, than the moments obtained from the less accurate observations.

4. Discussion

4.1. Interpretation of results

The current study addressed the question whether a more accurate assessment of occupational low-back load is worth the effort when it comes to accuracy and its predictive value for LBP. We conclude that although present at an individual level, differences between a moderately accurate and a higher accurate low-back load assessment methods are attenuated at a group level. For peak moments, differences between the two methods up to 300 Nm can be seen at an individual level. Such differences can be expected when using postural data in a biomechanical model (de Looze et al., 1994), and are mainly caused by inaccuracies in the observation and its rough categorization. Our analyses confirm this inaccuracy and add to this knowledge that errors might be proportional (i.e., differences between methods become proportionally larger with larger absolute moments). Also, differences between the two measurement methods decreased when moments were averaged over a group (mean systematic differences of about 40 Nm and 30 Nm for individual and group-based estimates). However, substantial differences remain present in this case. For cumulative moments, differences between the two estimation methods were limited and similar for individual level and group level estimates (i.e., showing comparable systematic and non-systematic differences between the two methods; Fig. 3). Correlations between the two methods were excellent, but slightly lower when an evidently outlying group was removed.

High within-subject variability as compared to between-subject variability in a pre-defined group of workers is common (Allread et al., 2000; Paquet et al., 2005), which attenuates the estimated LBP risk (i.e., OR) towards one (Mathiassen et al., 2010; Tielemans et al., 1998). Therefore, group-based approaches are often used in studies on exposure-outcome associations (e.g., Burdorf et al., 2006; Heederik et al., 2000) and have been shown to be an effective way to deal with this within-subject variability (Houba et al., 1997; Lyles et al., 1997). However, this advantage comes at the price of an increased uncertainty of the OR and thus a reduced power (Armstrong, 1998; Tielemans et al., 1998). This trade-off is supported by our results as it turns out that, although one can gain from using a measurement method with substantially higher accuracy at an individual level, such measurements do not necessarily lead to a better prediction of the outcome (i.e., LBP). This is evident in the comparable ORs obtained from the two methods (1.26 and 1.31), describing the association between peak low-back moments and LBP. Also for cumulative moments, ORs were comparable (ORs of 6.41 and 7.25, respectively) although our more accurate video-analysis method provided slightly higher ORs. Especially the cumulative low-back moments suggest substantial risks of LBP in the

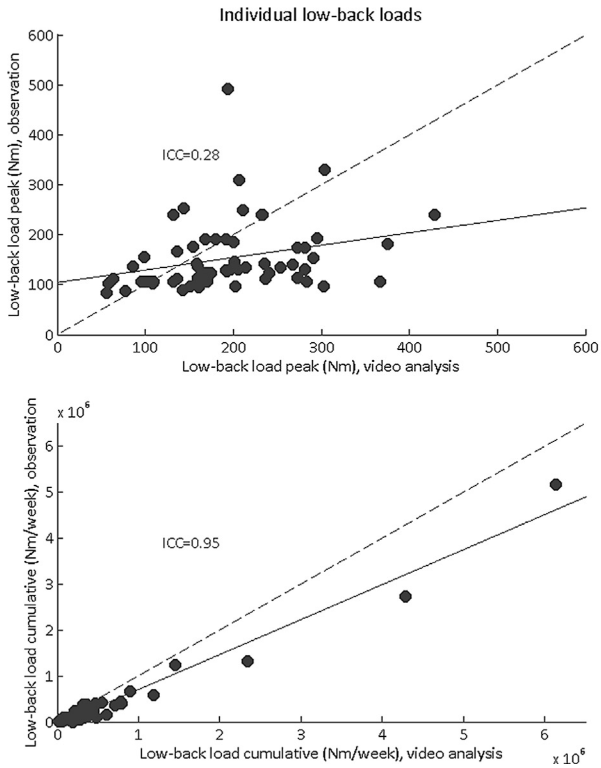

Fig. 2. Scatter plot depicting the association of low-back moments as obtained from structured observational data used in a biomechanical model (y-axis) and moments obtained from the video-analysis method (x-axis). Peak low-back moments (upper panels) and cumulative low-back moments (lower panels) at an individual level (left panels) and at a group level (right panels) are shown. The best fit through the data points (solid line) as well as the x = y reference line (dashed line) are shown. Also, intra-class correlation coefficients are shown, depicting the correlation between the two methods of calculation of moments.

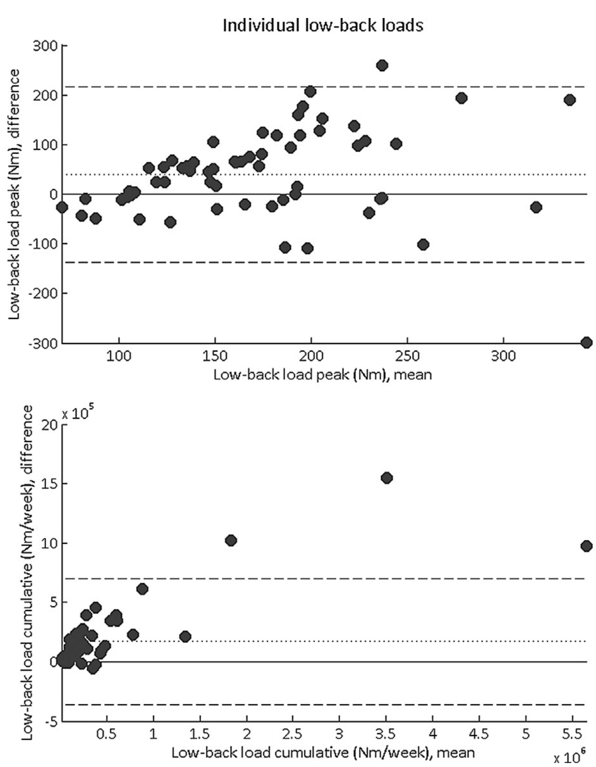

Fig. 3. Bland–Altman plot depicting the measurement differences (errors) between the two assessment methods plotted against their means. Peak low-back moments (upper panels) and cumulative low-back moments (lower panels) at an individual level (left panels) and at a group level (right panels) are shown. Dots represent the individual data-points, the dotted line represents the mean of all data points (systematic difference) and the dashed lines represent the 95% confidence interval of all the data-points (depicting the level of random differences between methods).

Table 3
Association of peak low-back moments and LBP for moments based on a biomechanical model using observational variables as input and moments obtained from a more accurate video-analysis method. Moments were obtained from 4872 MMH tasks from a representative sample of 93 workers. Mean moments with standard deviations (std) expressed in Nm are shown for the group of workers who reported LBP during the three years of follow up and for the group of workers who did not report LBP. ORs, 95% CIs and levels of significance are shown. In the right column, ORs corresponding with a difference of the groups with the highest low-back moment compared with the group and the lowest low-back moment are shown.

Metric	Method	Expressed in	LBP (n = 499)	No LBP (n = 632)	OR (95% CI)[a]	p-value	OR (high vs low)
			Mean load and std	Mean load and std			
Peak moment	Observation	Nm	86.39 (73.18)	82.20 (72.77)	1.001 (0.999–1.003)	0.211	1.26
	Video	Nm	110.13 (88.36)	104.83 (87.16)	1.001 (1.000–1.002)	0.139	1.31
Cumulative moment	Observation	10^5 Nm	2.67 (6.40)	1.76 (3.44)	1.056 (1.024–1.090)	0.001	6.41
	Video	10^5 Nm	1.66 (4.56)	0.98 (2.40)	1.086 (1.036–1.138)	0.001	7.25

OR = Odds Ratio, CI = confidence interval.
[a] Of 1131 workers data on the occurrence of LBP during follow-up and low-back loads were available.

group of workers with the highest mechanical loads (mainly road workers with high and repetitive external forces).

Our findings seem to be counter-intuitive as assessment methods that are more accurate are assumed to have a higher predictive power when assessing exposure-outcome associations (Burdorf, 2010). An important reason for this potentially is the high number of measurements per worker, which especially during integration in the cumulative moment calculation, reduces random errors and thus decreases differences in accuracy. This may be facilitated by the fact that subjects with high moments usually have a high number of MMH tasks. Therefore, provided that substantial numbers of MMH tasks and sufficient subjects are measured, a moderately accurate method to assess low-back loads can be sufficiently accurate to be significantly predictive for LBP.

Considering the above, it can be questioned whether a large investment (in terms of money and time) for measurements of low-back load is worth the effort. It has been shown earlier that the accuracy of a group-based exposure estimates improves progressively less when more group-members are included in the estimate (Hoozemans et al., 2001; Mathiassen et al., 2005). In line with this, the power of an exposure-outcome study improves progressively less as well, when measuring exposure from more workers within each group (Coenen et al., 2014b). Based on these outcomes as well as the present results, a large investment in research time and money does not necessarily lead to a proportional increase in accuracy of a measurement or study power. Our results can thus aid decision-making when designing new studies on exposure-outcome associations. However, more information (e.g. on the unit cost for obtaining exposure and outcome information and feasibility of the two alternative measurement methods) is required to make more informed decisions on the choice of a measurement tool. Therefore, a-priori pilot measurements, in which the accuracy, costs and feasibility of an exposure assessment method are evaluated, can guide researchers in making an educated decision on this matter. In addition, it should be taken into account that the cost of accurate methods may drastically decrease in the future, e.g. by use of inertial sensors or automated marker-less posture tracking methods (Dutta, 2012; Faber et al., 2010).

4.2. Methodological considerations

In the current study, we compared a rather inaccurate categorized posture assessment method of low-back loads to a relatively more accurate video-analysis method. Inaccuracy in the categorized posture observation method derives among other factors from a static moment calculation, categorized observed postures and the lack of categorization for some postures (e.g., lateral trunk bending). Differences between the two methods might stem from these issues and further improvements of the observation method might lead to even smaller differences between methods. The accuracy of our video-analysis method has been reported on before, showing good validity (i.e., high correspondence to a gold standard method; Coenen et al., 2011) and inter-rater reliability (i.e., high correspondance and low variation among raters; Coenen et al., 2013a). In spite of this higher accuracy, rather small, non-systematic, random errors for peak and mean low-back moments can still be expected (Coenen et al., 2011, 2013a). As there is no golden standard for physical load assessment in the field, we cannot speculate on whether even more accurate methods would lead to different outcomes.

A source of bias might stem from the fact that we only used two kinds of low-back load metrics (i.e., peak and cumulative moments). Although also other metrics are imaginable (e.g., measures of frequency or rest periods), we believe that we present a representative sample of metrics, that are often used in studies on the aetiology of LBP (Norman et al., 1998; Seidler et al., 2009). Besides, as an indicator of low-back load, resultant low-back moments were used. It is however known that also uniaxial loads, such as asymmetric low-back loads (occurring during occupational tasks) are important factors in the aetiology of LBP (Hoogendoorn et al., 2000). Results of our study are therefore limited in their representation of occupational work load. Moreover, it may be argued that injury risk and thus LBP is more accurately predicted by spinal forces, either in compression (van Dieën et al., 1999) or shear direction (Marras et al., 2010b; Norman et al., 1998). However, as a strong correlation of resultant low-back moments with shear forces and compression forces has been reported (van Dieën and Kingma, 2005), it is not expected that the choice of the metric has resulted in a higher bias in either of the two methods.

In our study, workers were recorded on video at four randomly chosen instants in one work day for a certain amount of time rather than during the whole work day. Distributing these four occasions over several days might have resulted in more precise low-back moment estimates, as work load will most likely vary more between days than within days (Mathiassen et al., 2003; Paquet et al., 2005). We addressed this issue by measuring several workers at different days in each group, to obtain more precise estimates of the work load within groups (Mathiassen et al., 2003; Paquet et al., 2005). Besides, this measurement strategy is justified by the small within group variability of exposure estimates in previous work on the SMASH cohort (Ariëns et al., 2001).

Moreover, in our study, low-back load was assessed from a limited number of four to five workers per group, introducing the possibility of selection bias. Such group-based approaches have been adopted before in work on the SMASH cohort (Ariëns et al., 2001; Hoogendoorn et al., 2000) and have proven to be successful in finding revealed several work-related physical risk factors of musculoskeletal disorders. Furthermore, a sensitivity analysis has shown that the current selection of workers is representative, leading to stable risk estimates for LBP (Coenen et al., 2014a). As mentioned above, in another study it has been shown that there is a *saturation* effect of increasing exposure sample sizes for OR bias and study power (Coenen et al., 2014b). In this study, bias and power decreased when exposure was collected from more workers up to the point where adding more subjects per group leads to only marginal improvements. More specifically, this study suggests that the OR depicting the exposure-outcome association did not show substantial attenuation after adding more than five workers per group. On the other hand, the size of the total population (i.e., of which LBP was assessed) showed to be of more importance for the outcome of an exposure-outcome association. We therefore consider the methods used in our study adequate.

On a related subject, the classification of workers into groups was made by the same trained observers who also collected the video recordings, on the basis of their training and experience in assessing physical workloads in occupational settings. In earlier research, grouping schemes have shown to influence the outcomes of a study (Symanski et al., 2006), for instance in terms of effectiveness in reducing attenuation of an exposure–outcome relationship (Werner and Attfield, 2000). Moreover, the categorisation of task groups was performed on a different dimension (i.e., generic work load) than the outcome of our study (i.e., low-back moments). Thus, another categorization of workers might have resulted in different outcomes. However, according to between group exposure contrasts, classification has been shown to be successful for the assessment of low-back loads in the same study population (Coenen et al., 2014b). Also we showed that the selection of workers for whom low-back moments were measured was highly comparable to the entire group of workers with respect to age, gender and prevalence of LBP (Table 1). Therefore, selection bias is not likely to have had a strong impact in the present study.

5. Conclusion

We studied whether a more accurate collection of occupational low-back load (which is often more time and money consuming) is worth the effort when it comes to accuracy and predictive value for LBP. From our results, it can be concluded that while estimates of peak low-back moments at the individual level can be improved substantially by using more accurate methods, this does not necessarily lead to a higher accuracy of low-back load estimates at a group level. A more accurate method also does not lead to better estimates of cumulative low-back moments. Although our study is prone to some limitations, our results can aid decision-making when designing new studies on low-back loads in specific tasks or groups, or on associations between low-back loads and LBP.

Conflicting interests

There are no conflicting interests.

Acknowledgements

No acknowledgements.

References

Allread, W.G., Marras, W.S., Burr, D.L., 2000. Measuring trunk motions in industry: variability due to task factors, individual differences, and the amount of data collected. Ergonomics 43, 691–701.

Ariëns, G.A., et al., 2001. Are neck flexion, neck rotation, and sitting at work risk factors for neck pain? Results of a prospective cohort study. Occup. Environ. Med. 58, 200–207.

Armstrong, B.G., 1998. Effect of measurement error on epidemiological studies of environmental and occupational exposures. Occup. Environ. Med. 55, 651–656.

Bakker, E.W., Verhagen, A.P., van Trijffel, E., Lucas, C., Koes, B.W., 2009. Spinal mechanical load as a risk factor for low back pain: a systematic review of prospective cohort studies. Spine 34, E281–E293.

Balogh, I., et al., 2004. Self-assessed and directly measured occupational physical activities – influence of musculoskeletal complaints, age and gender. Appl. Ergon. 35, 49–56.

Bland, J.M., Altman, D.G., 1986. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1, 307–310.

Burdorf, A., 2010. The role of assessment of biomechanical exposure at the workplace in the prevention of musculoskeletal disorders. Scand. J. Work Environ. Health 36, 1–2.

Burdorf, A., Jansen, J.P., 2006. Predicting the long term course of low back pain and its consequences for sickness absence and associated work disability. Occup. Environ. Med. 63, 522–529.

Chaffin, D.B., 2009. The evolving role of biomechanics in prevention of overexertion injuries. Ergonomics 52, 3–14.

Chang, C.C., Hsiang, S., Dempsey, P.G., McGorry, R.W., 2003. A computerized video coding system for biomechanical analysis of lifting tasks. Int. J. Ind. Ergon. 32, 239–250.

Coenen, P., Kingma, I., Boot, C.R., Bongers, P.M., Dieën van, J.H., 2014a. Cumulative mechanical low-back load at work is a determinant of low-back pain. Occup. Environ. Med. 71, 332–337.

Coenen, P., et al., 2011. Estimation of low back moments from video analysis: a validation study. J. Biomech. 44, 2369–2375.

Coenen, P., Kingma, I., Boot, C.R.L., Bongers, P.M., van Dieën, J.H., 2013a. Inter-rater reliability of a video-analysis method measuring low-back load in a field situation. Appl. Ergon. 44, 828–834.

Coenen, P., Kingma, I., Boot, C.R.L., Twisk, J.W.R., Bongers, P.M., van Dieën, J.H., 2013b. Cumulative low back load at work as a risk factor of low back pain: a prospective cohort study. J. Occup. Rehabil. 23, 11–18.

Coenen, P., Mathiassen, S.E., Kingma, I., Boot, C.R., Bongers, P.M., Dieën van, J.H., 2014b. Bias and power in group-based epidemiologic studies of low-back pain exposure and outcome – effects of study size and exposure measurement efforts. Ann. Occup. Hyg. 59, 439–454.

da Costa, B.R., Vieira, E.R., 2010. Risk factors for work-related musculoskeletal disorders: a systematic review of recent longitudinal studies. Am. J. Ind. Med. 53, 285–323.

David, G.C., 2005. Ergonomic methods for assessing exposure to risk factors for work-related musculoskeletal disorders. Occup. Med. 55, 190–199.

de Looze, M.P., Toussaint, H.M., Ensink, J., Mangnus, C., van der Beek, A.J., 1994. The validity of visual observation to assess posture in a laboratory-simulated, manual material handling task. Ergonomics 37, 1335–1343.

Dutta, T., 2012. Evaluation of the Kinect sensor for 3-D kinematic measurement in the workplace. Appl. Ergon. 43, 645–649.

Faber, G.S., Kingma, I., van Dieën, J.H., 2010. Bottom-up estimation of joint moments during manual lifting using orientation sensors instead of position sensors. J. Biomech. 43, 1432–1436.

Fleiss, J.L., 1986. The Design and Analysis of Clinical Experiments. John Wiley and Sons, Toronto.

Griffith, L.E., et al., 2012. Individual participant data meta-analysis of mechanical workplace risk factors and low back pain. Am. J. Public Health 102, 309–318.

Heederik, D., Attfield, M., 2000. Characterization of dust exposure for the study of chronic occupational lung disease: a comparison of different exposure assessment strategies. Am. J. Epidemiol. 151, 982–990.

Hoogendoorn, W.E., et al., 2000. Flexion and rotation of the trunk and lifting at work are risk factors for low back pain: results of a prospective cohort study. Spine 25, 3087–3092.

Hoozemans, M.J., Burdorf, A., van der Beek, A.J., Frings-Dresen, M.H., Mathiassen, S.E., 2001. Group-based measurement strategies in exposure assessment explored by bootstrapping. Scand. J. Work Environ. Health 27, 125–132.

Houba, R., Heederik, D., Kromhout, H., 1997. Grouping strategies for exposure to inhalable dust, wheat allergens and alpha-amylase allergens in bakeries. Ann. Occup. Hyg. 41, 287–296.

Jansen, J.P., Burdorf, A., 2003. Effects of measurement strategy and statistical analysis on dose-response relations between physical workload and low back pain. Occup. Environ. Med. 60, 942–947.

Kingma, I., Faber, G.S., van Dieën, J.H., 2010. How to lift a box that is too large to fit between the knees. Ergonomics 53, 1228–1238.

Kuorinka, I., et al., 1987. Standardised Nordic questionnaires for the analysis of musculoskeletal symptoms. Appl. Ergon. 18, 233–237.

Kwon, B.K., Roffey, D.M., Bishop, P.B., Dagenais, S., Wai, E.K., 2011. Systematic review: occupational physical activity and low back pain. Occup. Med. 61, 541–548.

Lötters, F., Burdorf, A., Kuiper, J., Miedema, H., 2003. Model for the work-relatedness of low-back pain. Scand. J. Work Environ. Health 29, 431–440.

Lyles, R.H., Kupper, L.L., Rappaport, S.M., 1997. A lognormal distribution-based exposure assessment method for unbalanced data. Ann. Occup. Hyg. 41, 63–76.

Marras, W.S., Lavender, S.A., Ferguson, S.A., Splittstoesser, R.E., Yang, G., 2010a. Quantitative biomechanical workplace exposure measures: distribution centers. J. Electromyogr. Kinesiol. 20, 813–822.

Marras, W.S., Lavender, S.A., Ferguson, S.A., Splittstoesser, R.E., Yang, G., 2010b. Quantitative dynamic measures of physical exposure predict low back functional impairment. Spine 35, 914–923.

Mathiassen, S.E., Burdorf, A., van der Beek, A.J., 2002. Statistical power and measurement allocation in ergonomic intervention studies assessing upper

trapezius EMG amplitude. A case study of assembly work. J. Electromyogr. Kinesiol. 12, 45–57.
Mathiassen, S.E., Moller, T., Forsman, M., 2003. Variability in mechanical exposure within and between individuals performing a highly constrained industrial work task. Ergonomics 46, 800–824.
Mathiassen, S.E., Nordander, C., Svendsen, S.W., Wellman, H.M., Dempsey, P.G., 2005. Task-based estimation of mechanical job exposure in occupational groups. Scand. J. Work Environ. Health 31, 138–151.
Mathiassen, S.E., Paquet, V., 2010. The ability of limited exposure sampling to detect effects of interventions that reduce the occurrence of pronounced trunk inclination. Appl. Ergon. 41, 295–304.
Neumann, W.P., Wells, R.P., Norman, R.W., Frank, J., Shannon, H., Kerr, M.S., 2001. A posture and load sampling approach to determining low-back pain risk in occupational settings. Int. J. Ind. Ergon. 27, 65–77.
Norman, R., Wells, R., Neumann, P., Frank, J., Shannon, H., Kerr, M., 1998. A comparison of peak vs cumulative physical work exposure risk factors for the reporting of low back pain in the automotive industry. Clin. Biomech. 13, 561–573.
Paquet, V., Punnett, L., Woskie, S., Buchholz, B., 2005. Reliable exposure assessment strategies for physical ergonomics stressors in construction and other non-routinized work. Ergonomics 48, 1200–1219.
Plamondon, A., Gagnon, M., Desjardins, P., 1996. Validation of two 3-D segment models to calculate the net reaction forces and moments at the L(5)/S(1) joint in lifting. Clin. Biomech. 11, 101–110.
Punnett, L., 2004. Work related neck pain: how important is it, and how should we understand its causes? Occup. Environ. Med. 61, 954–955.
Seidler, A., et al., 2009. Cumulative occupational lumbar load and lumbar disc disease – results of a German multi-center case-control study (EPILIFT). BMC Musculoskelet. Disord. 10, 48.
Symanski, E., Maberti, S., Chan, W., 2006. A meta-analytic approach for characterizing the within-worker and between-worker sources of variation in occupational exposure. Ann. Occup. Hyg. 50, 343–357.
Tielemans, E., Kupper, L.L., Kromhout, H., Heederik, D., Houba, R., 1998. Individual-based and group-based occupational exposure assessment: some equations to evaluate different strategies. Ann. Occup. Hyg. 42, 115–119.
Trask, C., et al., 2007. Measuring low back injury risk factors in challenging work environments: an evaluation of cost and feasibility. Am. J. Ind. Med. 50, 687–696.
van Dieën, J.H., Kingma, I., 2005. Effects of antagonistic co-contraction on differences between electromyography based and optimization based estimates of spinal forces. Ergonomics 48, 411–426.
van Dieën, J.H., Weinans, H., Toussaint, H.M., 1999. Fractures of the lumbar vertebral endplate in the etiology of low back pain: a hypothesis on the causative role of spinal compression in a specific low back pain. Med. Hypotheses 53, 246–252.
Wells, R., van Eerd, D., Hägg, G., 2004. Mechanical exposure concepts using force as the agent. Scand. J. Work Environ. Health 30, 179–190.
Werner, M.A., Attfield, M.D., 2000. Effect of different grouping strategies in developing estimates of personal exposures: specificity versus precision. Appl. Occup. Environ. Hyg. 15, 21–25.
Xu, X., Chang, C.C., Faber, G.S., Kingma, I., Dennerlein, J.T., 2010. Comparing polynomial and cubic spline interpolation of segment angles for estimating L5/S1 net moment during symmetric lifting tasks. J. Biomech. 43, 583–586.
Xu, X., Chang, C.C., Faber, G.S., Kingma, I., Dennerlein, J.T., 2011. The validity and inter-rater reliability of video-based posture observation during asymmetric lifting tasks. Hum. Factors 53, 371–382.
Zatsiorsky, V.M., 2002. Kinetics of Human Motion. Human Kinetics, Champaign, IL.